2) P ACC ACC TGG ACC ATC GCT GCA GAT GGT GGC AAG GCC TGA ATT

Construction of variant L351C+M430C+T183*+G184*: This variant was constructed by combining the L351C+M430C pairwise substitution mutation and the T183*+G184* pairwise deletion mutation by subdloning an approximately 1430 bp HindIII-Af/III fragment containing L351C+M430C into a pTVB106-like plasmid (with the T183*+G184* mutations) digested with the same enzymes.

Construction of variant Y243F+T183*+G184*: This variant was constructed by combining the Y243F mutation and the T183*+G184* mutation by subcloning an approximately 1148 bp DrdI fragment containing T183*+G184* into a pTVB106-like plasmid (with the Y243 mutation) digested with the same enzyme.

Bacillus subtilis transformants were screened for α-amylase activity on starch-containing agar plates and the presence of the correct mutations was checked by DNA sequencing.

Construction of variant Y243F+T183*+G184*+L351C+M430C: The L351C+M430C pairwise substitution mutation was subcloned as an approximately 470 bp XmaI-Sa/I fragment into a pTVB106-like vector (containing Y243F+T183*+G184*) digested with the same enzymes.

Construction of variant Y243F+T183*+G184*+L351C+M430C+Q391 E+K444Q: A pPM103-like vector containing the mutations Y243F+T183*+G184*+L351C+M430C was constructed by substituting the truncated version of SF16 in pPM103 with the approximately 1440 bp BstB1-Sa/I fragment of the pTVB106-like vector containing the five mutations in question. The Q391E and K444Q mutations were introduced simultaneously into the pPM103-like vector (containing Y243F+T183*+G184*+L351C+M430C) by the use of the following two mutagenesis primers in a manner similar to the previously described mutagenesis on pPM103:

P GGC AAA AGT TTG ACG TGC CTC GAG AA AGG GTC TAT

P TTG TCC CGC TTT ATT CTG GCC AA ATA CAT CCA TTT

B: Construction of Variants of the Parent α-amylase having the Amino Acid Sequence Shown in SEQ ID No. 2

Description of plasmid pTVB112: A vector, denoted pTVB112, to be used for the expression in *B. subtilis* of the α-amylase having the amino acid sequence shown in SEQ ID No. 2 was constructed. This vector is very similar to pTVB106 except that the gene encoding the mature α-amylase of SEQ ID No. 2 is inserted between the PstI and the HindIII sites in pTVB106. Thus, the expression of this α-amylase (SEQ ID No. 2) is also directed by the amyL promoter and signal sequence. The plasmid pTVB112 is shown in FIG. 4.

Construction of variant D183*+G184*: The construction of this variant was achieved using the PCR overlap extension mutagenesis method referred to earlier (vide supra). Primers #8573 and B1 were used in PCR reaction A, and primers #8569 and #8570 were used in PCR reaction B. The purified fragments from reaction A and reaction B and primers 1B and #8570 were used in PCR reaction C, resulting in an approximately 1020 bp DNA fragment. This fragment was digested with restriction endonucleases PstI and MIuI, and subcloned into the expression vector and transformed into *B. subtilis*.

Construction of further variants: By analogy with the construction (vide supra) of the plasmid pPM103 used in the production of mutants of the amino acid sequence of SEQ ID No. 1, a plasmid (denoted pTVB114; shown in FIG. 5) was constructed for the continued mutagenesis on variant D183*+G184*(SEQ ID No. 2). Mutations were introduced in pTVB114 (SEQ ID No. 2; D183*+G184*) in a manner similar to that for pPM103 (SEQ ID No. 1).

For the construction of the pairwise deletion variants R181*+D183* and R181*+G182*, it was chosen to alter the flanking amino acids in the variant D183*+G184* instead of deleting the specified amino acids in the wild type gene for SEQ ID No. 2. The following mutagenesis primer was used for the mutagenesis with pTVB114 as template:

PCC CAA TCC CAA GCT TTA CCA (T/C)CG AA TTG TAG ATA CG

The presence of a mixture of two bases (T/C) at one position allows for the presence of two different deletion flanking amino acid based on one mutagenesis primer. DNA sequencing of the resulting plasmids verifies the presence of either the one or the other mutation. The mutated gene of interest is subcloned as a PstI-DraIII fragment into pTVB112 digested with the same enzymes and transformed into *B. subtilis*.

For the construction of G182*+G184* and R181*+G184*, the following mutagenesis primer was used with pTVB114 as template:

PCC CAA TCC CAA GCT TTA TCT C(C/G)G AAC TTG TAG ATA CG

As before, the presence of a mixture of two bases (C/G) atone position allows for the presence of two different deletion flanking amino acid based on one mutagenesis primer. DNA sequencing of the resulting plasmids verifies the presence of either the one or the other mutation. The mutated gene of interest is subcloned as a PstI-DraIII fragment into pTVB112 digested with the same enzymes and transformed into *B. subtilis*.

For the construction of D183*+G184*+M202L the following mutagenesis primer was used:

PGA TCC ATA TCG ACG TCT GCA TAC AGT AAA TAA TC

For the construction of D183*+G184*+M202I the following mutagenesis primer was used:

PGA TCC ATA TCG ACG TCT GCA TAA ATT AAA TAA TC

EXAMPLE 3

Determination of Oxidation Stability of M202 Substitution Variants of the Parent α-amylases having the Amino Acid Sequences Shown in SEQ ID No. 1 and SEQ ID No. 2

A: Oxidation Stability of Variants of the Sequence in SEQ ID No. 1

The measurements were made using solutions of the respective variants in 50 mM Brifton-Robinson buffer (50 mM acetic acid, 50 mM phosphoric acid, 50 mM boric acid, 0.1 mM $CaCl_2$, pH adjusted to the value of interest with NaOH), pH 9.0, to which hydrogen peroxide was added (at time t=0) to give a final concentration of 200 mM H_2O_2. The solutions were then incubated at 40□C in a water bath.

After incubation for 5, 10, 15 and 20 minutes after addition of hydrogen peroxide, the residual α-amylase activity was measured using the Phadebas assay described above. The residual activity in the samples was measured using 50 mM Britton-Robinson buffer, pH 7.3, at 37□C (see Novo analytical publication AF207-1/1, available on request from Novo Nordisk A/S). The decline in activity was measured relative to a corresponding reference solution of the same enzyme at 0 minutes which was not incubated with hydrogen peroxide (100% activity).

The percentage of initial activity as a function of time is shown in the table below for the parent enzyme (SEQ ID No. 1) and for the variants in question.

	% Activity after incubation for (minutes)				
Variant	0	5	10	15	20
M202L	100	90	72	58	27
M202F	100	100	87	71	43
M202A	100	99	82	64	30
M202I	100	91	75	59	28
M2O2T	100	87	65	49	20
M202V	100	100	87	74	43
M202S	100	100	85	68	34
Parent	100	51	26	13	2

All the M202 substitution variants tested clearly exhibit significantly improved stability towards oxidation relative to the parent α-amylase (SEQ ID No. 1).

B: Oxidation Stability of Variants of the Sequence in SEQ ID No. 2

Measurements were made as described above using the parent α-amylase in question (SEQ ID No. 2), the variant M202L+D183*+G184*(designated L in the table below) and the variant M202I+D183*+G184*(designated I in the table below), respectively. In this case, incubation times (after addition of hydrogen peroxide) of 5, 10, 15 and 30 minutes were employed. As in the table above, the percentage of initial activity as a function of time is shown in the table below for the parent enzyme and for the variants in question.

	% Activity after incubation for (minutes)				
Variant	0	5	10	15	30
L	100	91	85	71	43
I	100	81	61	44	18
Parent	100	56	26	14	4

The two "substitution+pairwise deletion" variants tested (which both comprise an M202 substitution) clearly exhibit significantly improved stability towards oxidation relative to the parent α-amylase (SEQ ID No. 2).

EXAMPLE 4

Determination of Thermal Stability of Variants of the Parent α-amylases having the Amino Acid Sequences Shown in SEQ ID No. 1 and SEQ ID No. 2

A: Thermal Stability of Pairwise Deletion Variants of the Sequence in SEQ ID No. 1

Measurements were made using solutions of the respective variants in 50 mM Britton-Robinson buffer (vide supra), pH 9.0. The solutions were incubated at 65□C in a water bath, and samples were withdrawn after incubation for the indicated periods of time. The residual α-amylase activity of each withdrawn sample was measured using the Phadebas assay, as described above. The decline in activity was measured relative to a corresponding reference solution of the same enzyme at 0 minutes which was not incubated (100% activity).

The percentage of initial activity as a function of time is shown in the table below for the parent enzyme (SEQ ID No. 1) and for the following pairwise deletion variants in question:

Variant 1: R181*+G182*

Variant 2: R181*+T183*

Variant 3: G182*+G184*

Variant 4: T183*+G184*

Variant 5: T183*+G184*+R124P

	% Activity after incubation for (minutes)						
Variant	0	5	10	15	30	45	60
1	100	81	66	49	24	14	8
2	100	80	53	39	17	8	3
3	100	64	40	28	10	4	2
4	100	64	43	34	20	8	5
5	100	78	73	66	57	47	38
Parent	100	13	2	0	0	0	0

It is apparent that all of the pairwise deletion variants tested exhibit significantly improved thermal stability relative to the parent α-amylase (SEQ ID No. 1), and that the thermal stability of Variant 5, which in addition to the pairwise deletion mutation of Variant 4 comprises the substitution R124P, is markedly higher than that of the other variants. Since calorimetric results for the substitution variant R124P (comprising only the substitution R124P) reveal an approximately 7□C thermostabilization thereof relative to the parent α-amylase, it appears that the thermostabilizing effects of the mutation R124P and the pairwise deletion, respectively, reinforce each other.

B: Thermal Stability of Pairwise Deletion Variants of the Sequence in SEQ ID No. 2

Corresponding measurements were made for the parent enzyme (SEQ ID No. 2) and for the following pairwise deletion variants:

Variant A: D183*+G184*

Variant B: R181*+G182*

Variant C: G182*+G184*

	% Activity after incubation for (minutes)				
Variant	0	5	10	15	30
A	100	87	71	63	30
B	100	113	85	76	58
C	100	99	76	62	34
Parent	100	72	55	44	18

Again, it is apparent that the pairwise deletion variants in question exhibit significantly improved thermal stability relative to the parent α-amylase (SEQ ID No. 2).

C: Thermal Stability of a Multi-combination Variant of the Sequence in SEQ ID No. 1

Corresponding comparative measurements were also made for the following variants of the amino acid sequence shown in SEQ ID No. 1:

Variant 4: T183*+G184*

Variant 6: L351C+M430C

Variant 7: Y243F

Variant 8: Q391E+K444Q

Variant 9: T183*+G184*+L351C+M430C+Y243F+Q391E+K444Q

Variant	% Activity after incubation for (minutes)				
	0	5	10	15	30
4	100	66	41	22	7
6	100	87	73	65	43
7	100	14	2	1	0
8	100	69	46	31	14
9	100	92	93	89	82

Again, it appears that the thermostabilizing effect of multiple mutations, each of which has a thermostabilizing effect, is—at least qualitatively—cumulative.

EXAMPLE 5

Calcium-binding Affinity of α-amylase Variants of the Invention

Unfolding of amylases by exposure to heat or to denaturants such as guanidine hydrochloride is accompanied by a decrease in fluorescence. Loss of calcium ions leads to unfolding, and the affinity of a series of α-amylases for calcium can be measured by fluorescence measurements before and after incubation of each α-amylase (e.g. at a concentration of 1 μg/ml) in a buffer (e.g. 50 mM HEPES, pH 7) with different concentrations of calcium (e.g. in the range of 1 μM–100 mM) or of EGTA (e.g. in the range of 1–1000 μM) [EGTA=1,2-di(2-aminoethoxy)ethane-N,N,N', N'-tetraacetic acid] for a sufficiently long period of time such as 22 hours at 55□C).

The measured fluorescence F is composed of contributions form the folded and unfolded forms of the enzyme. The following equation can be derived to describe the dependence of F on calcium concentration ([Ca]):

$$F=[Ca]/(K_{diss}+[Ca])(\alpha_N-\beta_N \log([Ca]))+K_{diss}/(K_{diss}+[Ca])(\alpha_U-\beta_U \log([Ca]))$$

where α_N is the fluorescence of the native (folded) form of the enzyme, β_N is the linear dependence of α_N on the logarithm of the calcium concentration (as observed experimentally), α_U is the fluorescence of the unfolded form and β_U is the linear dependence of α_U on the logarithm of the calcium concentration. K_{diss} is the apparent calcium-binding constant for an equilibrium process as follows:

K_{diss}

N–Ca □ U+Ca (N=native enzyme; U=unfolded enzyme)

In fact, unfolding proceeds extremely slowly and is irreversible. The rate of unfolding is a dependent on calcium concentration, and the dependency for a given α-amylase provides a measure of the Ca-binding affinity of the enzyme. By defining a standard set of reaction conditions (e.g. 22 hours at 55□C), a meaningful comparison of K_{diss} for different α-amylases can be made. The calcium dissociation curves for α-amylases in general can be fitted to the equation above, allowing determination of the corresponding values of K_{diss}.

The following values for K_{diss} were obtained for the parent α-amylases having the amino acid sequences shown in SEQ ID No. 1 and SEQ ID No. 2, and for the indicated α-amylase variants according to the invention (the parent α-amylase being indicated in parentheses):

Variant	K_{diss} (mol/l)
D183* + G184* (SEQ ID No. 2)	$1.2 (\pm 0.5) \times 10^{-4}$
L351C + M430C + T183* + G184* (SEQ ID No. 1)	$1.7 (\pm 0.5) \times 10^{-3}$
T183* + G184* (SEQ ID No. 1)	$4.3 (\pm 0.7) \times 10^{-3}$
SEQ ID No. 2 (parent)	$4.2 (\pm 1.2) \times 10^{-2}$
SEQ ID No. 1 (parent)	$3.5 (\pm 1.1) \times 10^{-1}$

It is apparent from the above that the calcium-binding affinity of the latter α-amylolytic enzymes decreases in a downward direction through the above table, i.e. that the pairwise deletion variant D183*+G184*(SEQ ID No. 2) binds calcium most strongly (i.e. has the lowest calcium dependency) whilst the parent α-amylase of SEQ ID No. 1 binds calcium least strongly (i.e. has the highest calcium dependency).

REFERENCES CITED IN THE SPECIFICATION

Suzuki et al., the Journal of Biological Chemistry, Vol. 264, No. 32, Issue of November 15, pp. 18933–18938 (1989).

Hudson et al., Practical Immunology, Third edition (1989), Blackwell Scientific Publications.

Lipman and Pearson (1985) *Science* 227, 1435.

Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor, 1989.

S. L. Beaucage and M. H. Caruthers, *Tetrahedron Letters* 22, 1981, pp. 1859–1869.

Malthes et al., *The EMBO J.* 3, 1984, pp. 801–805.

R. K. Saiki et al., *Science* 239, 1988, pp. 487–491.

Morinaga et al., 1984, *Biotechnology* 2, pp. 646–639.

Nelson and Long, *Analytical Biochemistry* 180, 1989, pp. 147–151.

Hunkaplller et al., 1984, *Nature* 310, pp. 105–111.

R. Higuchi, B. Krummel, and R. K. Saiki (1988). A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucl. Acids Res.* 16, pp. 7351–7367.

Dubnau et al., 1971, *J. Mol. Biol.* 56, pp. 209–221.

Gryczan et al., 1978, *J. Bacteriol.* 134, pp. 318–329.

S. D. Erlich, 1977, *Proc. Natl. Acad. Sci.* 74, pp. 1680–1682.

Boel et al., 1990, *Biochemistry* 29, pp. 6244-6249.

Deng and Nickoloff, 1992, *Anal. Biochem.* 200, pp. 81–88.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(iii) NUMBER OF SEQUENCES: 32

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 485 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
His His Asn Gly Thr Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr
1               5                  10                  15

Leu Pro Asn Asp Gly Asn His Trp Asn Arg Leu Arg Asp Asp Ala Ala
            20                  25                  30

Asn Leu Lys Ser Lys Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Trp
        35                  40                  45

Lys Gly Thr Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr
    50                  55                  60

Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly
65                  70                  75                  80

Thr Arg Asn Gln Leu Gln Ala Ala Val Thr Ser Leu Lys Asn Asn Gly
                85                  90                  95

Ile Gln Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp
            100                 105                 110

Gly Thr Glu Ile Val Asn Ala Val Glu Val Asn Arg Ser Asn Arg Asn
        115                 120                 125

Gln Glu Thr Ser Gly Glu Tyr Ala Ile Glu Ala Trp Thr Lys Phe Asp
    130                 135                 140

Phe Pro Gly Arg Gly Asn Asn His Ser Ser Phe Lys Trp Arg Trp Tyr
145                 150                 155                 160

His Phe Asp Gly Thr Asp Trp Asp Gln Ser Arg Gln Leu Gln Asn Lys
                165                 170                 175

Ile Tyr Lys Phe Arg Gly Thr Gly Lys Ala Trp Asp Trp Glu Val Asp
            180                 185                 190

Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Met
        195                 200                 205

Asp His Pro Glu Val Ile His Glu Leu Arg Asn Trp Gly Val Trp Tyr
    210                 215                 220

Thr Asn Thr Leu Asn Leu Asp Gly Phe Arg Ile Asp Ala Val Lys His
225                 230                 235                 240

Ile Lys Tyr Ser Phe Thr Arg Asp Trp Leu Thr His Val Arg Asn Thr
                245                 250                 255

Thr Gly Lys Pro Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu
            260                 265                 270

Gly Ala Ile Glu Asn Tyr Leu Asn Lys Thr Ser Trp Asn His Ser Val
        275                 280                 285

Phe Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Asn Ser Gly
    290                 295                 300

Gly Tyr Tyr Asp Met Arg Asn Ile Leu Asn Gly Ser Val Val Gln Lys
305                 310                 315                 320
```

-continued

```
His Pro Thr His Ala Val Thr Phe Val Asp Asn His Asp Ser Gln Pro
                325                 330                 335

Gly Glu Ala Leu Glu Ser Phe Val Gln Gln Trp Phe Lys Pro Leu Ala
            340                 345                 350

Tyr Ala Leu Val Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr
        355                 360                 365

Gly Asp Tyr Tyr Gly Ile Pro Thr His Gly Val Pro Ala Met Lys Ser
    370                 375                 380

Lys Ile Asp Pro Leu Leu Gln Ala Arg Gln Thr Phe Ala Tyr Gly Thr
385                 390                 395                 400

Gln His Asp Tyr Phe Asp His His Asp Ile Ile Gly Trp Thr Arg Glu
                405                 410                 415

Gly Asn Ser Ser His Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asp
            420                 425                 430

Gly Pro Gly Gly Asn Lys Trp Met Tyr Val Gly Lys Asn Lys Ala Gly
        435                 440                 445

Gln Val Trp Arg Asp Ile Thr Gly Asn Arg Thr Gly Thr Val Thr Ile
    450                 455                 460

Asn Ala Asp Gly Trp Gly Asn Phe Ser Val Asn Gly Gly Ser Val Ser
465                 470                 475                 480

Val Trp Val Lys Gln
                485
```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 485 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
His His Asn Gly Thr Asn Gly Thr Met Met Gln Tyr Phe Glu Trp His
1               5                   10                  15

Leu Pro Asn Asp Gly Asn His Trp Asn Arg Leu Arg Asp Asp Ala Ser
            20                  25                  30

Asn Leu Arg Asn Arg Gly Ile Thr Ala Ile Trp Ile Pro Pro Ala Trp
        35                  40                  45

Lys Gly Thr Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr
    50                  55                  60

Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly
65                  70                  75                  80

Thr Arg Ser Gln Leu Glu Ser Ala Ile His Ala Leu Lys Asn Asn Gly
                85                  90                  95

Val Gln Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp
            100                 105                 110

Ala Thr Glu Asn Val Leu Ala Val Glu Val Asn Pro Asn Asn Arg Asn
        115                 120                 125

Gln Glu Ile Ser Gly Asp Tyr Thr Ile Glu Ala Trp Thr Lys Phe Asp
    130                 135                 140

Phe Pro Gly Arg Gly Asn Thr Tyr Ser Asp Phe Lys Trp Arg Trp Tyr
145                 150                 155                 160

His Phe Asp Gly Val Asp Trp Asp Gln Ser Arg Gln Phe Gln Asn Arg
                165                 170                 175
```

-continued

```
Ile Tyr Lys Phe Arg Gly Asp Gly Lys Ala Trp Asp Trp Glu Val Asp
            180                 185                 190

Ser Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Met
        195                 200                 205

Asp His Pro Glu Val Val Asn Glu Leu Arg Arg Trp Gly Glu Trp Tyr
    210                 215                 220

Thr Asn Thr Leu Asn Leu Asp Gly Phe Arg Ile Asp Ala Val Lys His
225                 230                 235                 240

Ile Lys Tyr Ser Phe Thr Arg Asp Trp Leu Thr His Val Arg Asn Ala
                245                 250                 255

Thr Gly Lys Glu Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu
            260                 265                 270

Gly Ala Leu Glu Asn Tyr Leu Asn Lys Thr Asn Trp Asn His Ser Val
        275                 280                 285

Phe Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Asn Ser Gly
    290                 295                 300

Gly Asn Tyr Asp Met Ala Lys Leu Leu Asn Gly Thr Val Val Gln Lys
305                 310                 315                 320

His Pro Met His Ala Val Thr Phe Val Asp Asn His Asp Ser Gln Pro
                325                 330                 335

Gly Glu Ser Leu Glu Ser Phe Val Gln Glu Trp Phe Lys Pro Leu Ala
            340                 345                 350

Tyr Ala Leu Ile Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr
        355                 360                 365

Gly Asp Tyr Tyr Gly Ile Pro Thr His Ser Val Pro Ala Met Lys Ala
    370                 375                 380

Lys Ile Asp Pro Ile Leu Glu Ala Arg Gln Asn Phe Ala Tyr Gly Thr
385                 390                 395                 400

Gln His Asp Tyr Phe Asp His His Asn Ile Ile Gly Trp Thr Arg Glu
                405                 410                 415

Gly Asn Thr Thr His Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asp
            420                 425                 430

Gly Pro Gly Gly Glu Lys Trp Met Tyr Val Gly Gln Asn Lys Ala Gly
        435                 440                 445

Gln Val Trp His Asp Ile Thr Gly Asn Lys Pro Gly Thr Val Thr Ile
    450                 455                 460

Asn Ala Asp Gly Trp Ala Asn Phe Ser Val Asn Gly Gly Ser Val Ser
465                 470                 475                 480

Ile Trp Val Lys Arg
                485
```

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 514 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
Ala Ala Pro Phe Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr Leu
1               5                   10                  15

Pro Asp Asp Gly Thr Leu Trp Thr Lys Val Ala Asn Glu Ala Asn Asn
            20                  25                  30
```

-continued

```
Leu Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys
        35                  40                  45

Gly Thr Ser Arg Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp
    50                  55                  60

Leu Gly Glu Phe Asn Gln Lys Gly Ala Val Arg Thr Lys Tyr Gly Thr
65                  70                  75                  80

Lys Ala Gln Tyr Leu Gln Ala Ile Gln Ala Ala His Ala Ala Gly Met
                85                  90                  95

Gln Val Tyr Ala Asp Val Val Phe Asp His Lys Gly Gly Ala Asp Gly
            100                 105                 110

Thr Glu Trp Val Asp Ala Val Glu Val Asn Pro Ser Asp Arg Asn Gln
        115                 120                 125

Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp Phe
    130                 135                 140

Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr His
145                 150                 155                 160

Phe Asp Gly Val Asp Trp Asp Glu Ser Arg Lys Leu Ser Arg Ile Tyr
                165                 170                 175

Lys Phe Arg Gly Ile Gly Lys Ala Trp Asp Trp Glu Val Asp Thr Glu
            180                 185                 190

Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met Asp His
        195                 200                 205

Pro Glu Val Val Thr Glu Leu Lys Ser Trp Gly Lys Trp Tyr Val Asn
    210                 215                 220

Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys
225                 230                 235                 240

Phe Ser Phe Phe Pro Asp Trp Leu Ser Asp Val Arg Ser Gln Thr Gly
                245                 250                 255

Lys Pro Leu Phe Thr Val Gly Glu Tyr Trp Ser Tyr Asp Ile Asn Lys
            260                 265                 270

Leu His Asn Tyr Ile Met Lys Thr Asn Gly Thr Met Ser Leu Phe Asp
        275                 280                 285

Ala Pro Leu His Asn Lys Phe Tyr Thr Ala Ser Lys Ser Gly Gly Thr
    290                 295                 300

Phe Asp Met Arg Thr Leu Met Thr Asn Thr Leu Met Lys Asp Gln Pro
305                 310                 315                 320

Thr Leu Ala Val Thr Phe Val Asp Asn His Asp Thr Glu Pro Gly Gln
                325                 330                 335

Ala Leu Gln Ser Trp Val Asp Pro Trp Phe Lys Pro Leu Ala Tyr Ala
            340                 345                 350

Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly Asp
        355                 360                 365

Tyr Tyr Gly Ile Pro Gln Tyr Asn Ile Pro Ser Leu Lys Ser Lys Ile
    370                 375                 380

Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln His
385                 390                 395                 400

Asp Tyr Leu Asp His Ser Asp Ile Ile Gly Trp Thr Arg Glu Gly Val
                405                 410                 415

Thr Glu Lys Pro Gly Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro
            420                 425                 430

Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Gln His Ala Gly Lys Val
        435                 440                 445
```

-continued

```
Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile Asn Ser
    450                 455                 460

Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser Val Trp
465                 470                 475                 480

Val Pro Arg Lys Thr Thr Val Ser Thr Ile Ala Trp Ser Ile Thr Thr
                485                 490                 495

Arg Pro Trp Thr Asp Glu Phe Val Arg Trp Thr Glu Pro Arg Leu Val
            500                 505                 510

Ala Trp
```

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1455 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```
CATCATAATG GAACAAATGG TACTATGATG CAATATTTCG AATGGTATTT GCCAAATGAC     60

GGGAATCATT GGAACAGGTT GAGGGATGAC GCAGCTAACT TAAAGAGTAA AGGGATAACA    120

GCTGTATGGA TCCCACCTGC ATGGAAGGGG ACTTCCCAGA ATGATGTAGG TTATGGAGCC    180

TATGATTTAT ATGATCTTGG AGAGTTTAAC CAGAAGGGGA CGGTTCGTAC AAAATATGGA    240

ACACGCAACC AGCTACAGGC TGCGGTGACC TCTTTAAAAA ATAACGGCAT TCAGGTATAT    300

GGTGATGTCG TCATGAATCA TAAAGGTGGA GCAGATGGTA CGGAAATTGT AAATGCGGTA    360

GAAGTGAATC GGAGCAACCG AAACCAGGAA ACCTCAGGAG AGTATGCAAT AGAAGCGTGG    420

ACAAAGTTTG ATTTTCCTGG AAGAGGAAAT AACCATTCCA GCTTTAAGTG GCGCTGGTAT    480

CATTTTGATG GGACAGATTG GGATCAGTCA CGCCAGCTTC AAAACAAAAT ATATAAATTC    540

AGGGGAACAG GCAAGGCCTG GGACTGGGAA GTCGATACAG AGAATGGCAA CTATGACTAT    600

CTTATGTATG CAGACGTGGA TATGGATCAC CCAGAAGTAA TACATGAACT TAGAAACTGG    660

GGAGTGTGGT ATACGAATAC ACTGAACCTT GATGGATTTA GAATAGATGC AGTGAAACAT    720

ATAAAATATA GCTTTACGAG AGATTGGCTT ACACATGTGC GTAACACCAC AGGTAAACCA    780

ATGTTTGCAG TGGCTGAGTT TTGGAAAAAT GACCTTGGTG CAATTGAAAA CTATTTGAAT    840

AAAACAAGTT GGAATCACTC GGTGTTTGAT GTTCCTCTCC ACTATAATTT GTACAATGCA    900

TCTAATAGCG GTGGTTATTA TGATATGAGA AATATTTTAA ATGGTTCTGT GGTGCAAAAA    960

CATCCAACAC ATGCCGTTAC TTTTGTTGAT AACCATGATT CTCAGCCCGG GGAAGCATTG   1020

GAATCCTTTG TTCAACAATG GTTTAAACCA CTTGCATATG CATTGGTTCT GACAAGGGAA   1080

CAAGGTTATC CTTCCGTATT TTATGGGGAT TACTACGGTA TCCCAACCCA TGGTGTTCCG   1140

GCTATGAAAT CTAAAATAGA CCCTCTTCTG CAGGCACGTC AAACTTTTGC CTATGGTACG   1200

CAGCATGATT ACTTTGATCA TCATGATATT ATCGGTTGGA CAAGAGAGGG AAATAGCTCC   1260

CATCCAAATT CAGGCCTTGC CACCATTATG TCAGATGGTC CAGGTGGTAA CAAATGGATG   1320

TATGTGGGGA AAAATAAAGC GGGACAAGTT TGGAGAGATA TTACCGGAAA TAGGACAGGC   1380

ACCGTCACAA TTAATGCAGA CGGATGGGGT AATTTCTCTG TTAATGGAGG GTCCGTTTCG   1440

GTTTGGGTGA AGCAA                                                    1455
```

(2) INFORMATION FOR SEQ ID NO: 5:

-continued

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1455 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CATCATAATG GGACAAATGG GACGATGATG CAATACTTTG AATGGCACTT GCCTAATGAT 60

GGGAATCACT GGAATAGATT AAGAGATGAT GCTAGTAATC TAAGAAATAG AGGTATAACC 120

GCTATTTGGA TTCCGCCTGC CTGGAAAGGG ACTTCGCAAA ATGATGTGGG GTATGGAGCC 180

TATGATCTTT ATGATTTAGG GGAATTTAAT CAAAAGGGGA CGGTTCGTAC TAAGTATGGG 240

ACACGTAGTC AATTGGAGTC TGCCATCCAT GCTTTAAAGA ATAATGGCGT TCAAGTTTAT 300

GGGGATGTAG TGATGAACCA TAAAGGAGGA GCTGATGCTA CAGAAAACGT TCTTGCTGTC 360

GAGGTGAATC CAAATAACCG GAATCAAGAA ATATCTGGGG ACTACACAAT TGAGGCTTGG 420

ACTAAGTTTG ATTTTCCAGG GAGGGGTAAT ACATACTCAG ACTTTAAATG GCGTTGGTAT 480

CATTTCGATG GTGTAGATTG GGATCAATCA CGACAATTCC AAAATCGTAT CTACAAATTC 540

CGAGGTGATG GTAAGGCATG GGATTGGGAA GTAGATTCGG AAAATGGAAA TTATGATTAT 600

TTAATGTATG CAGATGTAGA TATGGATCAT CCGGAGGTAG TAAATGAGCT TAGAAGATGG 660

GGAGAATGGT ATACAAATAC ATTAAATCTT GATGGATTTA GGATCGATGC GGTGAAGCAT 720

ATTAAATATA GCTTTACACG TGATTGGTTG ACCCATGTAA GAAACGCAAC GGGAAAAGAA 780

ATGTTTGCTG TTGCTGAATT TTGGAAAAAT GATTTAGGTG CCTTGGAGAA CTATTTAAAT 840

AAAACAAACT GGAATCATTC TGTCTTTGAT GTCCCCCTTC ATTATAATCT TTATAACGCG 900

TCAAATAGTG GAGGCAACTA TGACATGGCA AAACTTCTTA ATGGAACGGT TGTTCAAAAG 960

CATCCAATGC ATGCCGTAAC TTTTGTGGAT AATCACGATT CTCAACCTGG GGAATCATTA 1020

GAATCATTTG TACAAGAATG GTTTAAGCCA CTTGCTTATG CGCTTATTTT AACAAGAGAA 1080

CAAGGCTATC CCTCTGTCTT CTATGGTGAC TACTATGGAA TTCCAACACA TAGTGTCCCA 1140

GCAATGAAAG CCAAGATTGA TCCAATCTTA GAGGCGCGTC AAAATTTTGC ATATGGAACA 1200

CAACATGATT ATTTTGACCA TCATAATATA ATCGGATGGA CACGTGAAGG AAATACCACG 1260

CATCCCAATT CAGGACTTGC GACTATCATG TCGGATGGGC CAGGGGGAGA GAAATGGATG 1320

TACGTAGGGC AAAATAAAGC AGGTCAAGTT TGGCATGACA TAACTGGAAA TAAACCAGGA 1380

ACAGTTACGA TCAATGCAGA TGGATGGGCT AATTTTTCAG TAAATGGAGG ATCTGTTTCC 1440

ATTTGGGTGA AACGA 1455

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1548 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GCCGCACCGT TTAACGGCAC CATGATGCAG TATTTTGAAT GGTACTTGCC GGATGATGGC 60

ACGTTATGGA CCAAAGTGGC CAATGAAGCC AACAACTTAT CCAGCCTTGG CATCACCGCT 120

CTTTGGCTGC CGCCCGCTTA CAAAGGAACA AGCCGCAGCG ACGTAGGGTA CGGAGTATAC 180

-continued

GACTTGTATG ACCTCGGCGA ATTCAATCAA AAAGGGACCG TCCGCACAAA ATACGGAACA 240

AAAGCTCAAT ATCTTCAAGC CATTCAAGCC GCCCACGCCG CTGGAATGCA AGTGTACGCC 300

GATGTCGTGT TCGACCATAA AGGCGGCGCT GACGGCACGG AATGGGTGGA CGCCGTCGAA 360

GTCAATCCGT CCGACCGCAA CCAAGAAATC TCGGGCACCT ATCAAATCCA AGCATGGACG 420

AAATTTGATT TTCCCGGGCG GGGCAACACC TACTCCAGCT TTAAGTGGCG CTGGTACCAT 480

TTTGACGGCG TTGATTGGGA CGAAAGCCGA AAATTGAGCC GCATTTACAA ATTCCGCGGC 540

ATCGGCAAAG CGTGGGATTG GGAAGTAGAC ACGGAAAACG GAAACTATGA CTACTTAATG 600

TATGCCGACC TTGATATGGA TCATCCCGAA GTCGTGACCG AGCTGAAAAA CTGGGGGAAA 660

TGGTATGTCA ACACAACGAA CATTGATGGG TTCCGGCTTG ATGCCGTCAA GCATATTAAG 720

TTCAGTTTTT TTCCTGATTG GTTGTCGTAT GTGCGTTCTC AGACTGGCAA GCCGCTATTT 780

ACCGTCGGGG AATATTGGAG CTATGACATC AACAAGTTGC ACAATTACAT TACGAAAACA 840

GACGGAACGA TGTCTTTGTT TGATGCCCCG TTACACAACA AATTTTATAC CGCTTCCAAA 900

TCAGGGGGCG CATTTGATAT GCGCACGTTA ATGACCAATA CTCTCATGAA AGATCAACCG 960

ACATTGGCCG TCACCTTCGT TGATAATCAT GACACCGAAC CCGGCCAAGC GCTGCAGTCA 1020

TGGGTCGACC CATGGTTCAA ACCGTTGGCT TACGCCTTTA TTCTAACTCG GCAGGAAGGA 1080

TACCCGTGCG TCTTTTATGG TGACTATTAT GGCATTCCAC AATATAACAT TCCTTCGCTG 1140

AAAAGCAAAA TCGATCCGCT CCTCATCGCG CGCAGGGATT ATGCTTACGG AACGCAACAT 1200

GATTATCTTG ATCACTCCGA CATCATCGGG TGGACAAGGG AAGGGGGCAC TGAAAAACCA 1260

GGATCCGGAC TGGCCGCACT GATCACCGAT GGGCCGGGAG GAAGCAAATG GATGTACGTT 1320

GGCAAACAAC ACGCTGGAAA AGTGTTCTAT GACCTTACCG GCAACCGGAG TGACACCGTC 1380

ACCATCAACA GTGATGGATG GGGGGAATTC AAAGTCAATG GCGGTTCGGT TTCGGTTTGG 1440

GTTCCTAGAA AAACGACCGT TTCTACCATC GCTCGGCCGA TCACAACCCG ACCGTGGACT 1500

GGTGAATTCG TCCGTTGGAC CGAACCACGG TTGGTGGCAT GGCCTTGA 1548

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 485 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

His His Asn Gly Thr Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr
1 5 10 15

Leu Pro Asn Asp Gly Asn His Trp Asn Arg Leu Asn Ser Asp Ala Ser
20 25 30

Asn Leu Lys Ser Lys Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Trp
35 40 45

Lys Gly Ala Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr
50 55 60

Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly
65 70 75 80

Thr Arg Ser Gln Leu Gln Ala Ala Val Thr Ser Leu Lys Asn Asn Gly
85 90 95

Ile Gln Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp

-continued

```
           100                 105                 110

Ala Thr Glu Met Val Arg Ala Val Glu Val Asn Pro Asn Asn Arg Asn
        115                 120                 125

Gln Glu Val Thr Gly Glu Tyr Thr Ile Glu Ala Trp Thr Arg Phe Asp
    130                 135                 140

Phe Pro Gly Arg Gly Asn Thr His Ser Ser Phe Lys Trp Arg Trp Tyr
145                 150                 155                 160

His Phe Asp Gly Val Asp Trp Asp Gln Ser Arg Arg Leu Asn Asn Arg
                165                 170                 175

Ile Tyr Lys Phe Arg Gly His Gly Lys Ala Trp Asp Trp Glu Val Asp
            180                 185                 190

Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Ile Asp Met
        195                 200                 205

Asp His Pro Glu Val Val Asn Glu Leu Arg Asn Trp Gly Val Trp Tyr
    210                 215                 220

Thr Asn Thr Leu Gly Leu Asp Gly Phe Arg Ile Asp Ala Val Lys His
225                 230                 235                 240

Ile Lys Tyr Ser Phe Thr Arg Asp Trp Ile Asn His Val Arg Ser Ala
                245                 250                 255

Thr Gly Lys Asn Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu
            260                 265                 270

Gly Ala Ile Glu Asn Tyr Leu Gln Lys Thr Asn Trp Asn His Ser Val
        275                 280                 285

Phe Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Lys Ser Gly
    290                 295                 300

Gly Asn Tyr Asp Met Arg Asn Ile Phe Asn Gly Thr Val Val Gln Arg
305                 310                 315                 320

His Pro Ser His Ala Val Thr Phe Val Asp Asn His Asp Ser Gln Pro
                325                 330                 335

Glu Glu Ala Leu Glu Ser Phe Val Glu Glu Trp Phe Lys Pro Leu Ala
            340                 345                 350

Tyr Ala Leu Thr Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr
        355                 360                 365

Gly Asp Tyr Tyr Gly Ile Pro Thr His Gly Val Pro Ala Met Arg Ser
    370                 375                 380

Lys Ile Asp Pro Ile Leu Glu Ala Arg Gln Lys Tyr Ala Tyr Gly Lys
385                 390                 395                 400

Gln Asn Asp Tyr Leu Asp His His Asn Ile Ile Gly Trp Thr Arg Glu
                405                 410                 415

Gly Asn Thr Ala His Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asp
            420                 425                 430

Gly Ala Gly Gly Ser Lys Trp Met Phe Val Gly Arg Asn Lys Ala Gly
        435                 440                 445

Gln Val Trp Ser Asp Ile Thr Gly Asn Arg Thr Gly Thr Val Thr Ile
    450                 455                 460

Asn Ala Asp Gly Trp Gly Asn Phe Ser Val Asn Gly Gly Ser Val Ser
465                 470                 475                 480

Ile Trp Val Asn Lys
                485
```

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs

-continued

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GCTGCGGTGA CCTCTTTAAA AAATAACGGC 30

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CCACCGCTAT TAGATGCATT GTAC 24

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CTTACGTATG CAGACGTCGA TATGGATCAC CC 32

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GATCCATATC GACGTCTGCA TACGTAAGAT AGTC 34

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

TTASGGGCAA GGCCTGGGAC TGG 23

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

CCCAGGCCTT GCCCSTAAAT TTATATATTT TGTTTTG 37

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid

-continued

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GGTTTCGGTT CGAAGGATTC ACTTCTACCG C 31

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GCGGTAGAAG TGAATCCTTC GAACCGAAAC CAG 33

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

GGTACTATCG TAACAATGGC CGATTGCTGA CGCTGTTATT TGC 43

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

CTGTGACTGG TGAGTACTCA ACCAAGTC 28

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

CTACTTCCCA ATCCCAAGCT TTACCTCGGA ATTTG 35

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

CAAATTCCGA GGTAAAGCTT GGGATTGGGA AGTAG 35

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

-continued

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

TTGAACAACC GTTCCATTAA GAAG 24

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

CTCTGTATCG ACTTCCCAGT CCCAAGCTTT TGTCCTGAAT TTATATATTT TGTTTTGAA 60

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

CTCTGTATCG ACTTCCCAGT CCCAAGCTTT GCCTCCGAAT TTATATATTT TGTTTTGAA 60

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

ATGTGTAAGC CAATCGCGAG TAAAGCTAAA TTTTATATGT TTCACTGCAT C 51

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

GCACCAAGGT CATTTCGCCA GAATTCAGCC ACTG 34

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

TGTCAGAACC AACGCGTATG CACATGGTTT AAACCATTG 39

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

-continued

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

CCACCTGGA CCATCGCTGC AGATGGTGGC AAGGCCTGAA TT 42

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

GGCAAAAGTT TGACGTGCCT CGAGAAGAGG GTCTAT 36

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

TTGTCCCGCT TTATTCTGGC CAACATACAT CCATTT 36

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

CCCAATCCCA AGCTTTACCA YCGAACTTGT AGATACG 37

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

CCCAATCCCA AGCTTTATCT CSGAACTTGT AGATACG 37

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

GATCCATATC GACGTCTGCA TACAGTAAAT AATC 34

(2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

-continued

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

GATCCATATC GACGTCTGCA TAAATTAAAT AATC 34

What is claimed is:

1. A variant of a parent *Bacillus stearothermophilus* alpha-amylase, wherein the variant has an amino acid sequence which has at least 95% homology to the parent *Bacillus stearothermophilus* alpha-amylase and comprises a deletion of amino acids 179 an 180, using SEQ ID NO:3 for numbering, and wherein the variant has alpha-amylase activity.

2. The variant of claim 1, wherein the variant further comprises a substitution of a cysteine at amino acids 349 and 428, using SEQ ID NO:3 for numbering.

3. A variant alpha-amylase, wherein the variant has at least 95% homology to SEQ ID NO:3 and comprises a deletion of amino acids 179 and 180, using SEQ ID NO:3 for numbering and wherein the variant has alpha-amylase activity.

4. The variant of claim 3, wherein the variant further comprises a substitution of a cysteine at amino acids 349 and 428, using SEQ ID NO:3 for numbering.

5. A variant of a *Bacillus stearothermophilus* alpha-amylase, wherein the alpha-amylase variant consists of a deletion of amino acids 179 and 180, using SEQ ID NO:3 for numbering.

* * * * *

EXHIBIT B

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

NOVOZYMES A/S, Plaintiff v. GENENCOR INTERNATIONAL, INC., and ENZYME DEVELOPMENT CORPORATION Defendants	C.A. No. 05-160-KAJ

DECLARATION OF FRANCES HAMILTON ARNOLD

I, Frances Hamilton Arnold, do hereby declare as follows:

1. I am a citizen of the United States, and am more than twenty-one (21) years of age.

2. I am the Dickinson Professor of Chemical Engineering and Biochemistry at the California Institute of Technology (CIT), where I have been employed since 1986. From 1985-1986, before I began working at CIT, I was a Research Fellow in the Department of Chemistry and the University of California, Berkeley.

3. In 1985, I earned a Ph.D. degree in Chemical Engineering from the University of California, Berkeley. I received a B.S. degree *magna cum laude* from Princeton University in 1979.

4. I have received several professional awards and honors. These include: Institute of Medicine of the National Academies (2004); National Academy of Engineering (2000); Food, Pharmaceuticals, and Bioengineering Division Award, American Institute of Chemical